F COPY

AFGL-TR-88-0116

AD-A203 542

TR-5216-1

EVALUATION OF THE AFGL CLOUD SIMULATION MODEL

Gary Rasmussen
Paul Janota
Ralph Ferraro

Research and Data Systems ~~Corp~~ Inc
10300 Greenbelt Road
Lanham, MD 20706

October 1987

Scientific Report No. 1

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

AIR FORCE GEOPHYSICS LABORATORY
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
HANSCOM AIR FORCE BASE, MASSACHUSETTS 01731-5000

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; Distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFGL-TR-88-0116
6a. NAME OF PERFORMING ORGANIZATION Research and Data Systems ~~Corporation~~ Inc 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Geophysics Laboratory
6c. ADDRESS (City, State, and ZIP Code) 10300 Greenbelt Road Lanham, MD 20706	7b. ADDRESS (City, State, and ZIP Code) Hanscom AFB Massachusetts 01731-5000
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F19628-85-C-0007
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO
62101F	6670	09	AN

11. TITLE (Include Security Classification)
Evaluation of the AFGL Cloud Simulation Model

12. PERSONAL AUTHOR(S)
Gary Rasmussen*; Paul Janota*; Ralph Ferraro

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Scientific Report #1	FROM ____ TO ____	1987 October	54

16. SUPPLEMENTARY NOTATION
* The Analytic Sciences Corp, 55 Walkers Brook Drive, Reading, MA 01867 (Cont fr plc)

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Sky cover distributions; Monte Carlo simulation; Cloud distribution; Cloud detection; Sample size determination; Sky cover database; Model validation

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This report documents progress in validating two empirical models of the distribution of total sky cover over lines and areas. The Burger Area Algorithm (BAA), as applied to sky cover, is a mathematical model of the probability that fraction C or less of area A is covered by clouds, given scale length r and mean clear p0 (one minus the mean sky cover). The Burger Line Algorithm (BLA) is similar except it gives the probability of linear coverage rather than areal coverage. The validation procedure adopted relies on sky cover distributions over lines and areas of various sizes at three U.S. sites representative of several homogeneous climatic regions. The distributions will be obtained empirically from satellite imagery. The six tasks ~~have been~~ identified are: 1)

Task 1 - Select Test Areas; 2)

Task 2 - Determine Sample Sizes; 3)

Task 3 - Acquire Satellite Imagery; 4)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Oliver Muldoon, Capt, USAF	22b. TELEPHONE (Include Area Code) (617) 377-5957 — 22c. OFFICE SYMBOL AFGL/LYA

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

CONT OF BLOCK 19:

Task 4 - Develop Cloud Detection and Analysis Algorithms

Task 5 - Create a Sky Cover Database

Task 6 - Determine Error Bounds for the BAA and BLA Models

This report discusses the output of Tasks 1 through 4 which are currently complete (April 1985). The methodology employed during Task 2 (Monte Carlo simulation of the validation process) is described in detail.

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION

The purpose of this project is to determine error bounds for two empirical models of total sky cover climatologies: The Burger Area Algorithm (BAA) and the Burger Line Algorithm (BLA). These models were developed by Charles Burger and Irving Gringorten at the U.S Air Force Geophysics Laboratory (AFGL) and are documented in their technical report (Ref. 1).

The major tasks to be completed during this project are discussed in Chapter 2. Briefly, they are:

- Task 1 - Select Test Areas
- Task 2 - Determine Sample Sizes
- Task 3 - Acquire Satellite Imagery
- Task 4 - Develop Cloud Detection and Analysis Software
- Task 5 - Create a Cloud Cover Database
- Task 6 - Determine Error Bounds For the AFGL Models

During the period of this report we have completed the first four tasks. Construction of the cloud cover database is currently underway.

The rest of this report is organized as follows. Chapter 2 provides a more detailed description of project goals and study tasks. In addition, results for the completed tasks are discussed. Chapter 3 contains a detailed discussion

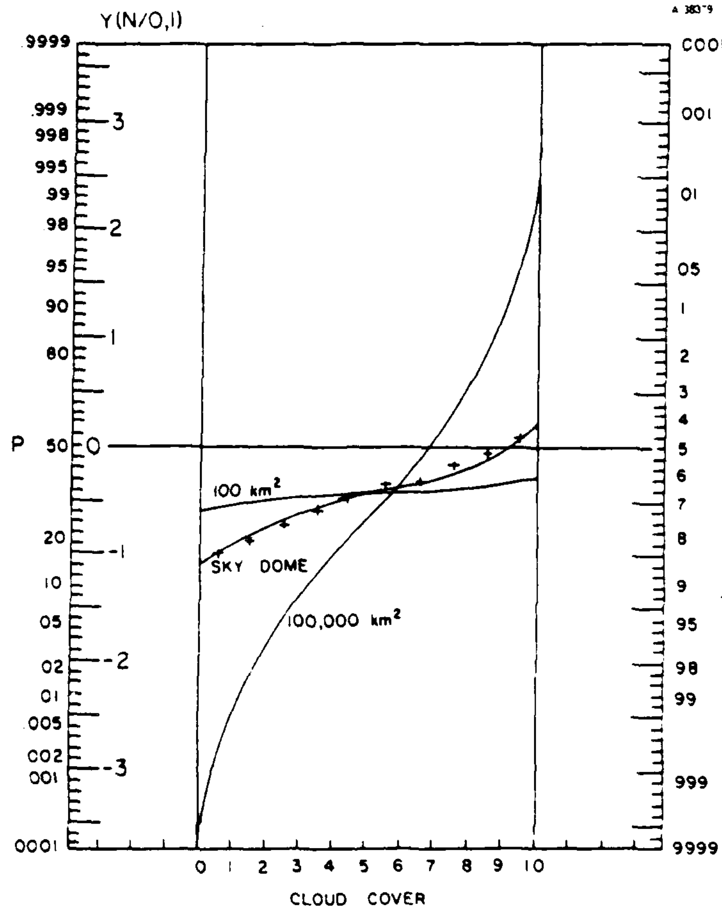
of the assumptions, methods, and results of the sample size study. Chapter 4 summarizes current progress and outlines our future plans. Appendix A describes the software algorithms developed to process the GOES imagery. Finally, Appendix B describes the organization of the image statistics files.

2. PROJECT OVERVIEW

2.1 PURPOSE OF THE PROJECT

RDS and its subcontractor, The Analytic Sciences Corporation (TASC), are validating the AFGL statistical cloud modeling algorithms known as the Burger Area Algorithm (BAA) and the Burger Line Algorithm (BLA). These models were developed at AFGL by Burger and Gringorten and are described in their Technical Report (Ref. 1) which discusses the model equations and applications to sky cover estimation. The BAA is a mathematical model of the probability pA that the areal fractional cloud cover cA is less than or equal to a threshold areal coverage CA, given the area A, the scale distance r, and a probability p0 that a point on the surface is not cloud covered. The scale distance r is a measure of the range of spatial autocorrelation of cloud cover. Similarly, the BLA is a mathematical model of the probability pL that the lineal fractional coverage cL is less than or equal to a threshold lineal coverage CL, given a line length L, scale distrance r, and probability p0 that a point on the surface is not cloud covered. Note that for a point on the surface p0 is equivalent to the mean clearness at a point and 1-p0 is the mean sky cover, a statistic commonly available from various climate records. The BAA and BLA are essentially curve fits to data samples generated by multiple runs of the Boehm Sawtooth Wave (SW) model (Ref. 1).

Figure 2.1-1 shows an example of the application of the BAA by comparing the cumulative cloud cover distributions for different reference areas derived from the same historical sky dome from data at Bedford, MA.



FROM BURGER, C.F. AND I.I GRINGORTEN, 1984: TWO-DIMENSIONAL MODELING FOR LINEAL AND AREAL PROBABILITIES OF WEATHER CONDITIONS. AFGL-TR-84-0126.

Figure 2.1-1 Total Cloud Cover at Bedford, Mass., January, 1200-1400 Local Time.

2.2 STUDY APPROACH

The RDS/TASC team is using bi-spectral GOES imagery to test the validity of the AFGL algorithms by empirically deriving relationships between cloud cover and reference areas and line lengths and comparing these to relationships predicted by the algorithms. In order to accomplish this objective, there are six major study tasks:

- Select test areas
- Determine the size of the image sample required for significance
- Acquire GOES visible and IR images
- Develop cloud detection and analysis programs
- Create a cloud cover database for reference lines and areas
- Generate AFGL model statistics and compare with satellite derived values.

During the period of this report we have completed the first four tasks and are in the process of creating the cloud cover database.

2.2.1 Selecting Test Areas

Test areas were selected based on the following criteria:

- Cloud cover distributions should be significantly different from one area to another
- Cloud climate in each area should be as spatially homogeneous as possible
- Horizontal distance from each area to the GOES subpoint should be as small as practical
- Each area should be centered on a surface reporting station associated with a RUSSWO[1] climate summary

[1] Revised Uniform Summary of Surface Weather Observations

- Surface observations coincident with the image samples should be available.

Based on a survey of available RUSSWO's and application of the other criteria, the study team proposed to investigate three areas centered on Ft. Riley, KS; Rickenbacker AFB, OH; and Key West, FL (the area around Key West is somewhat offset in order to minimize inhomogeneities created by the Florida mainland). Each area is approximately 316 x 316km. Surface reports in the USAF/AWS DATSAV format are available coincident with the image sample period.

2.2.2 Determining Image Sample Size

The procedure for determining an acceptable image sample size is the subject of Chapter 3 of this report. Based on this procedure, it was shown that a satisfactory validation could be performed in each area using sample sizes of 450 or more images for each time of day and season. Consequently, we decided to use five years of bi-spectral imagery acquired twice per day during the winter and summer months from 1979 through 1983 for a grand total of 5460 subscenes.

2.2.3 Acquisition of GOES Data

The visible and IR subscenes were acquired from NESDIS in digital form in the highest resolution and gray scale available from the archives. Visible images have [illegible] gray shades and a resolution of about 1km; IR images have [illegible] gray shades and a resolution of about 8km.

2.2.4 Developing Automated Cloud Detection and Analysis Programs

The RDS/TASC team developed an automated bi-spectral procedure to detect clouds on GOES imagery, compare cloud cover with observed ground truth, reject suspect images for manual interactive evaluation, and extract areal and lineal database parameters. Appendix A summarizes the logic of this process. Initial tests of the program have shown that approximately 20 percent of the imagery fails the initial quality control checks and must be manually analyzed.

All phases of the procedure are run on TASC's Alliant FX-8 mini supercomputer using a modified version of the TASC Interactive Image Processing System (TIIPS). The processing of each image and insertion of required statistical values into the database requires 24 seconds per image on the Alliant FX-8.

2.2.5 Creation of the Cloud Cover Database

This task will begin during the final contract phase. This involves running all of the 5460 subscenes through the database software programs at TASC and building a dataset in a format suitable for processing by AFGL. The process will complete datasets in station sequence so that AFGL can begin their phase of the work during the database build task. The contents of the delivered dataset files are shown in Appendix B.

2.2.6 Generation of Model Statistics and Comparison with Satellite-Derived Values

AFGL will take the satellite-based data set values of p0 and r and will create model-based distributions of cloud cover as functions of area and line length and will return these data

to the contractor for final validation of the AFGL algorithms. RDS, assisted by TASC, will compare the model-based distributions with the image-based distributions using a family of significance tests including the Komolgorov-Smirnov (K-S) test, the log-likelihood ratio (G) test, and the chi-square test (see Ref. 5). Chapter 3 expands on the validation process in considerable additional detail.

3. DETERMINING SAMPLE SIZES FOR GOODNESS-OF-FIT TESTS

3.1 INTRODUCTION

The purpose of this chapter is to summarize progress made in estimating sample sizes required in the main project task of estimating error bounds for two mathematical models of total sky cover climatologies: the Burger Area Algorithm (BAA) and the Burger Line Algorithm (BLA). These models were developed by Charles Burger and Irving Gringorten at the U.S. Air Force Geophysics Laboratory (AFGL). Their technical report (Ref. 1) discusses the model equations and application of the models to total sky cover.

The BAA is a mathematical model of the probability pA that the areal fractional coverage cA is less than or equal to a threshold areal coverage CA, given area A, scale distance r (a measure of the range of spatial autocorrelation), and probability p0 that a point is not covered:

$$pA = \text{Prob}\ [\ cA <= CA \mid A,\ r,\ p0\ ] \qquad (3.1\text{-}1)$$

When CA is one we must interpret pA as the probability that the areal coverage cA is strictly less than CA. The BLA is similar. It is a mathematical model of the probability pL that the lineal fractional coverage cL, is less than or equal to a threshold lineal coverage CL given line length L, scale distance r, and probability p0:

$$pL = \text{Prob}\ [\ cL <= CL \mid L,\ r,\ p0\ ] \qquad (3.1\text{-}2)$$

Again, if CL is one, then we must interpret pL as the probability that the lineal coverage cL is strictly less than CL. Note that the BAA and BLA models assume homogeneous and stationary field statistics. Note also that pA and pL are cumulative probabilities; therefore, BAA and BLA are (respectively) areal and lineal coverage cumulative distribution functions (CDFs).

Generally, a point is covered if the field value at that point exceeds a predetermined threshold value. When applied to total sky cover, a convenient threshold is zero for vertically integrated cloud water content. That is, a surface point is covered if cloud water exists anywhere along the geocentric ray passing through the surface point. Otherwise it is not covered. Note that p0 is the probability that a point is <u>not</u> covered. Thus, when applied to total sky cover, p0 is the mean clear (1-mean sky cover). Fraction cA of an area (or cL of a line) is covered when fraction cA (or cL) of the set of constituent points are covered.

3.2 APPLICATION CONSIDERATIONS

The AFGL BLA and BAA models are applicable to a wide variety of scalar, two-dimensional, stochastic fields. They give easy to use, easy to compute CDFs for coverage over areas and along lines for a wide variety of weather parameters. Two examples of practical applications could be:

- determining the probability of a cloud-free line of sight along a trajectory, and
- determining the probability of a half inch or more of precipitation in 24 hours over 80 percent of the area of a drainage basin.

It is reasonable to expect the magnitude of the model error bounds to be application dependent. In general the error bounds should depend on the amount of deviation between actual conditions experienced in an application and the ideal conditions which were imposed during model development.

An important source of deviation between ideal and actual conditions is the spatial autocorrelation function. The correlation function built into the models is that of Boehm's Sawtooth Wave Model (Ref.1, pp 10-14). It has a single fixed shape, although the scale distance parameter is available to account for varying ranges of correlation. Natural phenomena, on the other hand, exhibit correlation structures of many shapes. It is not surprising then that analytic correlation functions of a variety of forms have been fitted to observational data and used in applications.

Four analytic correlation functions commonly used in meteorology are plotted in Figs. 3.2-1a through 3.2-4a. The Sawtooth Wave correlation function is plotted in Fig. 3.2-5a. Note that each function is isotropic and homogeneous. While such assumptions are convenient and are frequently used, natural correlation structures are often significantly anisotropic. Moreover, they are approximately homogeneous only in sufficiently small spatial domains.

Comparison of the shapes of the plotted correlation functions has been facilitated by standardizing each so that the correlation at zero separation is one and the correlation at unit separation is the reciprocal of e, the base of the natural logarithms. All five functions agree at two points; elsewhere they are free to differ. In nature, as in the figures, they may differ substantially.

A-47183

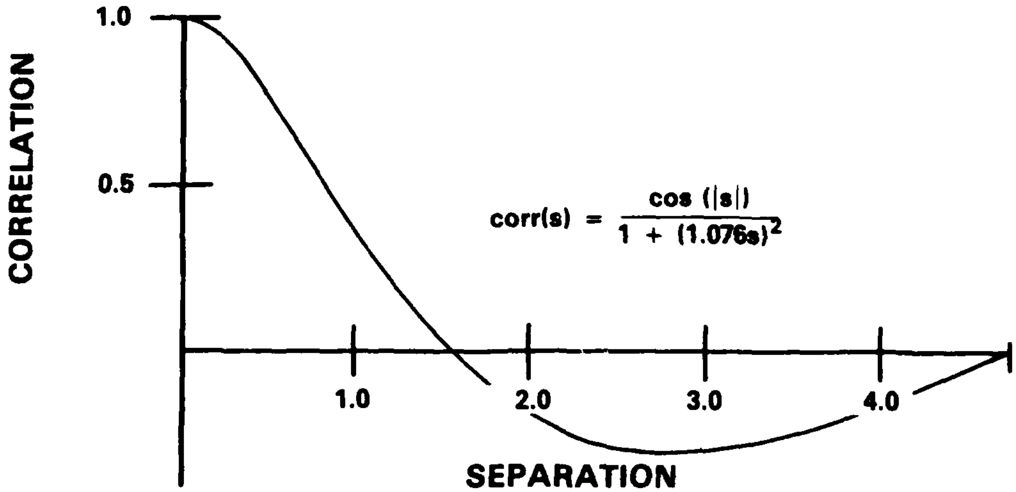


Figure 3.2-1a Damped Cosine Correlation Function

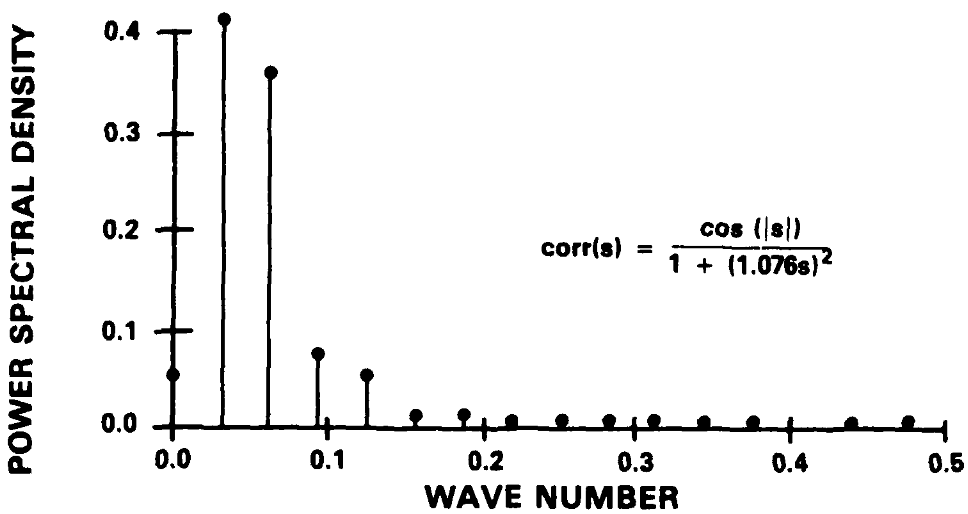


Figure 3.2-1b Damped Cosine Periodogram

A-47184

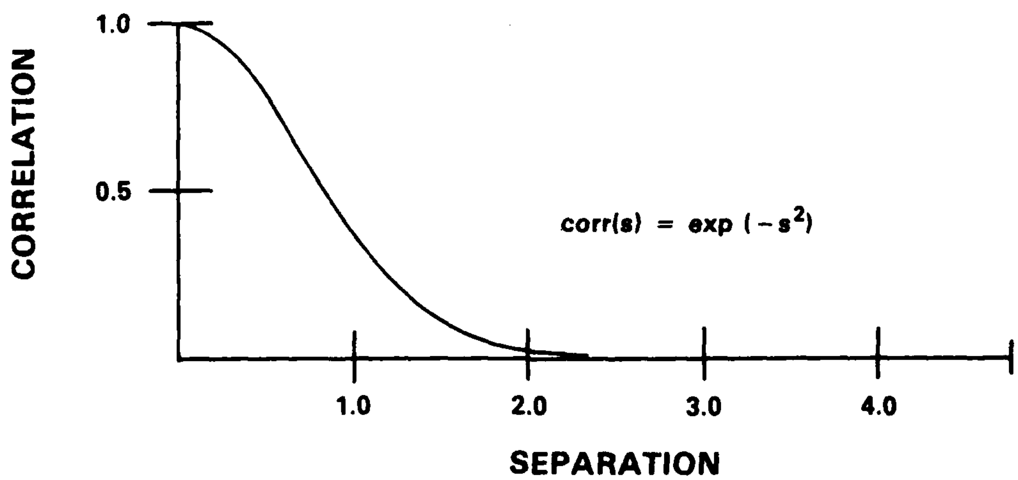


Figure 3.2-2a Gaussian Correlation Function

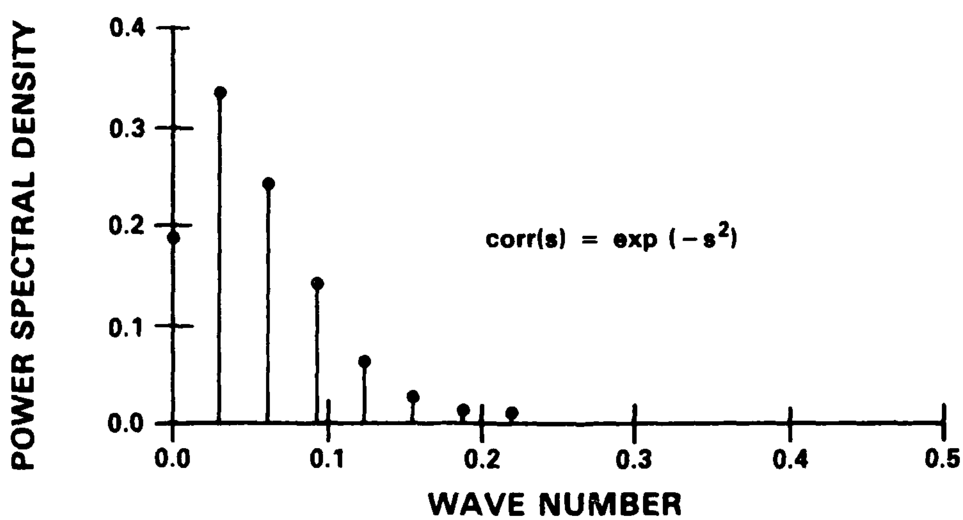


Figure 3.2-2b Gaussian Periodogram

A-47185

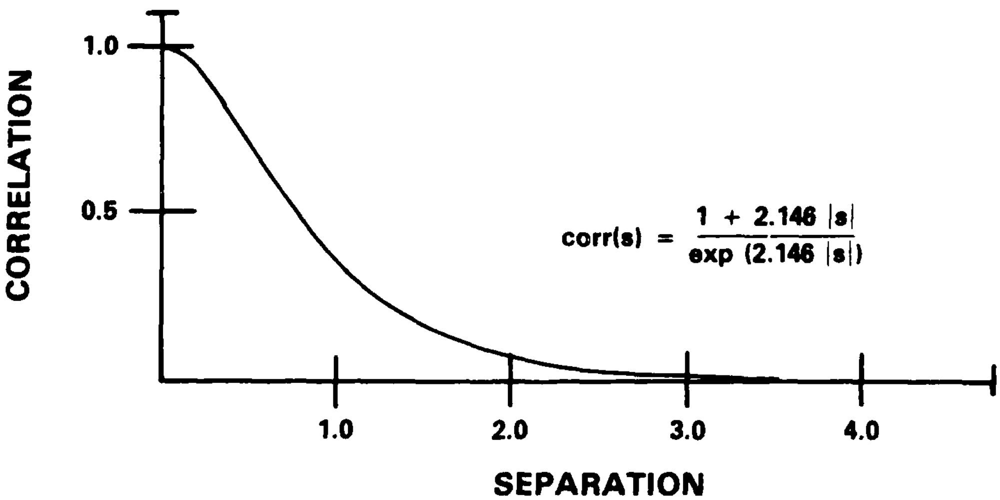


Figure 3.2-3a SOAR Correlation Function

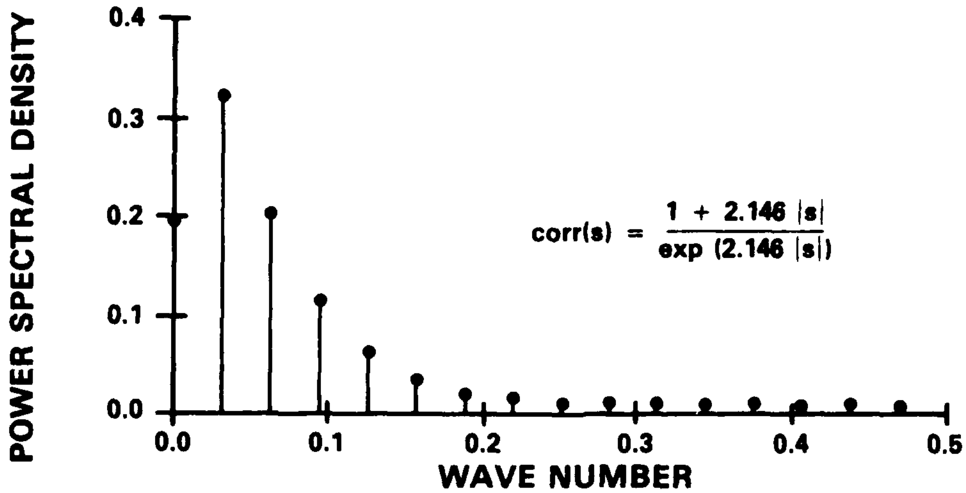


Figure 3.2-3b SOAR Periodogram

A-47186

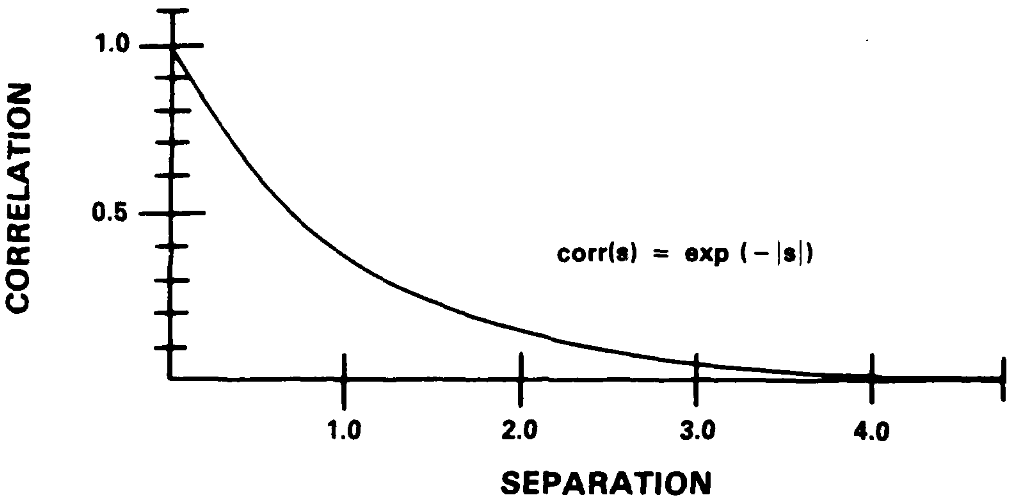


Figure 3.2-4a Exponential Correlation Function

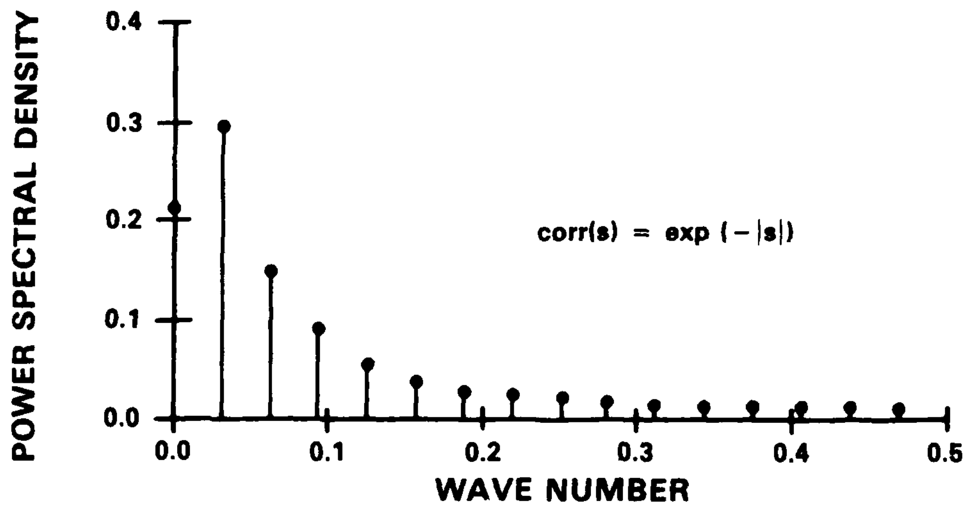


Figure 3.2-4b Exponential Periodogram

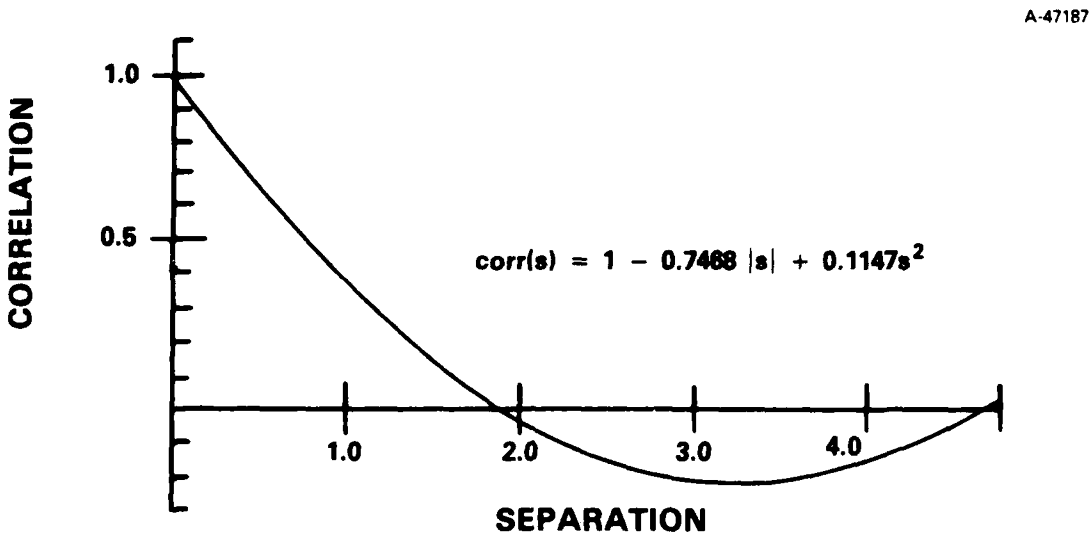


Figure 3.2-5a Boehm SW Correlation Function

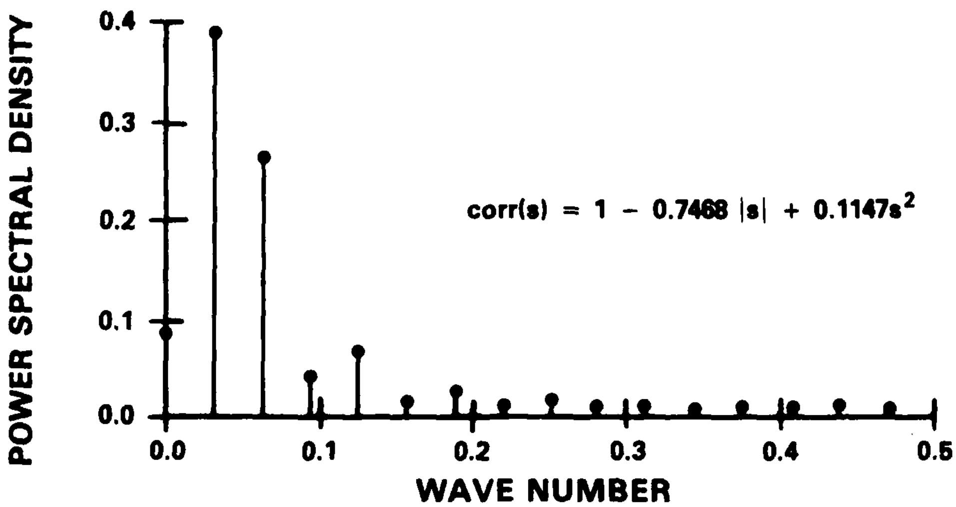


Figure 3.2-5b Boehm SW Periodogram

An alternate view of each correlation function is provided by its Fourier transform (see Figs. 3.2-1b through 3.2-5b.) The Fourier transform of an autocorrelation function is also a periodogram for the underlying variate. That is, it gives the fraction of the total mean square variation in the underlying variate which may be attributed to each spectral band. The periodograms in the figures have been standardized to give a total mean square variation of one. Comparing Figs. 3.2-1 through 3.2-4 with Fig. 3.2-5 provides us with a clue about how the model error bounds may vary as the models are applied to actual meteorological variates. We infer that the AFGL model error bounds should be relatively large when applied to a weather parameter such as geopotential which is characterized by a correlation function similar to those in Figs. 3.2-1 through 3.2-3. Similarly, the bounds should be relatively small when applied to a weather parameter such as surface relative humidity which is characterized by a correlation function similar to the exponential function in Fig. 3.2-4.

3.3 PROJECT GOALS AND SAMPLE SIZES

The AFGL models were developed by fitting curves to CDFs derived from an analysis of a very large synthetic data set. While the models agree well with the development data, the extent of agreement with observational data must be determined for each application. The goal of this project is to determine error bounds for the models as applied to satellite observations of total sky cover over lines and areas.

The basic project plan is as follows. Empirical CDFs for total sky cover will be constructed by analyzing a series of imagery from a Geostationary Operational Environmental Satellite (GOES). Statistical tests of goodness-of-fit between the empirical and corresponding theoretical (model predicted) CDFs will

then determine the model error bounds. Employing GOES data for three widely separated sites, several seasons, and several times of day and determining sky cover for several sizes of lines and areas will enable error bounds to be estimated over a wide range of values of model parameters.

Prior to ordering GOES imagery, it was necessary to determine the sample size required to estimate error bounds with acceptable reliability (defined below). About five years of imagery were available. Using three months of data per season, per site, per time of day would yield a sample size of about 450 (5y x 3m/y x 30d/m). Using one month of data per season would reduce the nominal sample size to about 150 but would enable examination of additional seasons and times of day. This is desirable as it will allow us to locate test points more densely in the models' parameter spaces. In short, we wanted to order the smallest samples which would enable us to reliably estimate model error bounds.

Because of sampling variation, small samples can be expected to result in relatively large differences between model and empirical CDFs, even when the model CDFs are perfectly accurate. Conversely, large samples can be expected to result in relatively small differences, providing that the model CDFs are correct. The developers of the BAA and BLA models estimated their absolute accuracy to be about 0.05. Thus, we needed a sample size which would give us reasonable assurance of detecting significant errors that small. Therefore, the required sample size was defined to be the smallest sample size which would permit detection, with 95 percent confidence, of significant model CDF errors as small as 0.05 in magnitude. In other words, if the absolute difference between an empirical cumulative probability and the corresponding model predicted cumulative probability is greater than 0.05, then we want to be able to conclude with 95

percent confidence that the difference is due to model failure rather than sampling variation. In this context we needed to know if samples of 150 were large enough or if samples of 300 (two months per season) or 450 (three months per season) were required. The rest of this chapter summarizes work done to answer these questions.

3.4 QUICK ESTIMATES OF THE REQUIRED SAMPLE SIZE

Several quick estimates of the required sample size can be obtained by employing approximations and simplifications. The first involves a general approach to sample size determination (Ref. 2, pp 516-517). The key idea is to focus on a single point in the CDF rather than the entire CDF.

Let X be a statistic with standard deviation sd(X) and estimated value e(X). Then, with confidence 100*(1-a), a two-sided confidence interval for X may be approximated by:

$$[e(X)-Z_{a/2}*sd(X),\ e(X)+Z_{a/2}*sd(X)] \qquad (3.4\text{-}1)$$

where Z_q is the upper qth percentile of the standard normal distribution. If we can tolerate an error in e(X) of magnitude L or smaller, then Eq. (3.4-1) can be rewritten:

$$sd(X) <= L/Z_{a/2} \qquad (3.4\text{-}2)$$

Assuming that sd(X) in Eq. (3.4-2) can be expressed as a function of sample size N, we can solve for N and get an estimate of the required sample size.

For our purposes, let X in Eq. (3.4-2) be the proportion P of the sample with sky cover less than or equal to a specified threshold (i.e., P is one point from the CDF). Thus P

is given by either pA (from the BAA) or pL (from the BLA). We know (Ref. 3, p 175) that the standard deviation of a proportion is given by:

$$sd(P) = [P*(1-P)/N]^{0.5} \tag{3.4-3}$$

Substituting this expression for sd(P) in Eq. (3.4-2) and solving for sample size N yields:

$$N >= P*(1-P)*(Z_{a/2}/L)^2 \tag{3.4-4}$$

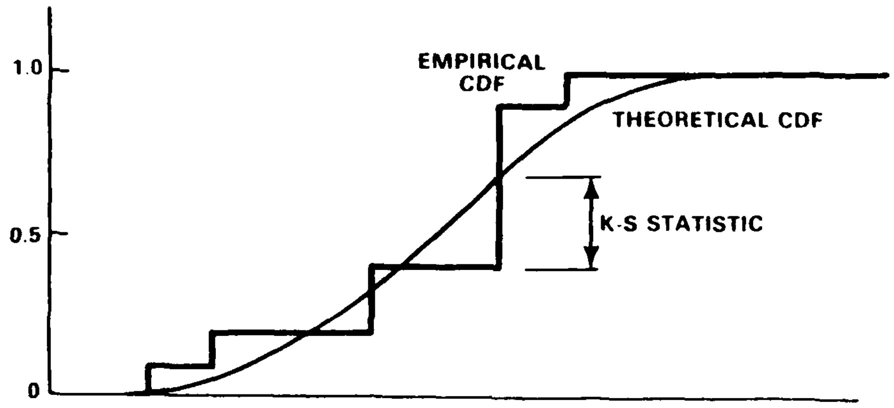
The worst case (largest N) occurs for P=0.5. This is plausible since the two ends of a CDF are tied to values of 0 and 1 while between these limits much variation is possible. The discussion in Section 3 above suggests L and a should both be set to 0.05. Using these values of P, L, and a ($Z_{0.025}$=1.96), yields a minimum required sample size of approximately 385. So, this analysis indicates that a sample size of 450 is adequate while sample sizes of 150 and 300 are not. We conclude that we will need to use data from all three months of each season considered to reduce random sampling errors down to the level of anticipated model errors.

A second way to get a quick estimate of the required sample size is to look at the large sample asymptotic behavior of the critical values of the Kolmogorov-Smirnov (K-S) goodness-of-fit statistic D. Figure 3.4-1 provides a quick overview of this test statistic. D is defined to be the greatest absolute difference between the theoretical and empirical CDFs and thus D is a parameter of direct interest and interpretation. For a sample size larger than 30 and the standard K-S test, an approximate 95 percent confidence critical value is given by (Ref. 4, p 203):

$$D_{0.05} \doteq 1.358\ N^{-0.5} <=L \tag{3.4-5}$$

A 38381

- DEFINED AS MAXIMUM ABSOLUTE DIFFERENCE BETWEEN THEORETICAL AND EMPIRICAL CUMULATIVE FREQUENCY DISTRIBUTIONS



- ASSUMES A CONTINUOUS VARIATE
- FOR EXTRINSIC HYPOTHESIS AND N > 30

$$D_{.05} = \frac{1.358}{\sqrt{N}} \qquad D = 0.05 \Rightarrow N = 738$$

- FOR INTRINSIC HYPOTHESIS AND N > 30

$$D_{0.05} = \frac{0.886}{\sqrt{N}} \qquad D = 0.05 \Rightarrow N = 314$$

Figure 3.4-1 The Kolmogorov-Smirnov Statistic

An estimate of the required sample size may be obtained by setting L equal to the error which can be tolerated in the CDF, in this case 0.05. Solving for N yields a minimum required sample size of about 738. This estimate is much larger than the previous one and is suspect because an important assumption of the K-S test is violated (Ref. 5, p 718).

The violated assumption, known as the extrinsic hypothesis, states that the parameters of the theoretical (model) distribution are known independently of the data at hand (i.e., they are not estimated from the same data used to compute the test statistic). The assumption is violated in this project because the same set of satellite scenes which are used to determine model parameters for each site, season, and time of day will be used to construct the empirical CDFs.

An alternate K-S test applicable with a normal populations and an intrinsic hypothesis (i.e., the same data are used to estimate distribution parameters and construct the empirical CDF) was developed by Lilliefors (Ref. 6). In his formulation, an approximate 95 percent confidence critical value is given by (Ref. 4, p 206):

$$D_{0.05} = 0.886\ N^{-0.5} <= L \qquad (3.4\text{-}6)$$

Again, an estimate of the required sample size may be obtained by setting L equal to 0.05. Solving for N yields a minimum required sample size of about 314. This is much closer to our first estimate of 385. We again conclude that samples of 150 and 300 are not adequate and that samples of 450 are adequate.

The assumption of a normal population in the preceding analysis is probably not valid. However, the actual shape of the underlying distribution probably would not impact our choice of

450 as sample size. Evidence for this assertion may be found in a paper by Crutcher (Ref. 7). He presents expressions for the large sample asymptotic critical values for the K-S statistic which are valid with an intrinsic hypothesis and data from a wide range of distribution shapes (exponential, gamma, normal, and extreme value). The expressions for the 95 percent confidence critical values are all similar to Eqs. (3.4-5) and (3.4-6). The only difference is in the constant which ranges between 0.886 for a normal population and 1.06 for a particular exponential population. The latter constant value yields the largest sample size, namely 449. So again we conclude that a sample of 450 is probably adequate while samples of 150 and 300 are not.

3.5 STATISTICAL GOODNESS-OF-FIT TESTS

The basic methodology for determining model error bounds in this project will be statistical goodness-of-fit between empirical and model-predicted (theoretical) statistical distributions. Three different goodness-of-fit tests, (Ref. 5, pp 691-731) will be employed. Two of the tests, the Kolmogorov-Smirnov (K-S) test and the G test, each have unique advantages. The third test, the chi-square test, will be included because it is the traditional goodness-of-fit test and because including it will not significantly impact project cost.

The K-S test is summarized in Fig. 3.4-1 and is described briefly in the previous section. With continuous data, it is the most powerful test of the three considered here. Moreover, as the greatest absolute difference between empirical and model predicted cumulative relative frequency distributions, and K-S test statistic is a parameter of direct interest.

Unfortunately, the standard K-S test cannot be used to validate the AFGL models because three of its assumptions will be violated:

- satellite observations of sky cover will be discrete rather than continuous
- model parameters will be estimated from the data (intrinsic null hypothesis) rather than being known <u>a priori</u>.
- sky cover data from sequential days will be serially correlated (linearly dependent) rather than independent.

The continuity assumption will be most nearly correct for large areas with many potential values of sky cover (many pixels per scene). It will be least valid for short lines with few potential values for sky cover. Violation of the continuity assumption makes the standard K-S test more conservative (Ref. 5, p 720). That is, the standard test will reject a false null hypothesis of equality of distributions less often than expected under the stated significance level. Several modifications of the standard K-S test to account for discrete data have been proposed (Refs. 8 and 9).

The assumption of an extrinsic null hypothesis is invalid. Two model parameters (scale distance and mean sky cover) will be estimated from the data for one size line or area and applied with all sizes of lines and areas. Violation of the extrinsic assumption makes the standard K-S test more conservative. As discussed above in Section 3.4, test modifications are available for some populations.

The degree of serial correlation in sky cover data will probably vary with location, time of day, and season. On average, however, the 24 hr serial correlation of sky cover is slightly greater than 0.2 (Ref. 10, p 8). Correlations of

this magnitude probably will not have a major impact on test results. In general, violation of the assumption of independence reduces the effective sample size and makes the standard K-S test less conservative.

None of the available K-S tests, standard or modified, can account for all of the violations of assumptions which will occur when determining error bounds for the AFGL models. Thus, we will not be able to employ published tables of critical values. Consequently, in order to use a K-S type of test, it will be necessary to estimate test statistic critical values through some other means such as Monte Carlo simulation of the entire validation process.

The G test statistic for goodness-of-fit is based on information theory. It is defined to be twice the amount of information in the sample which is available for discriminating between the expected distribution and the observed distribution:

$$G = 2\ [O_1 \ln (O_1/E_1) + \ldots + O_k \ln (O_k/E_k)] \qquad (3.5\text{-}1)$$

where k is the number of cells, O_i is the observed frequency in cell i, and E_i is the expected (model predicted) frequency in cell i. Test statistic G is approximately distributed as a chi-square variate with k-1-p degrees of freedom where p is the number of distribution parameters which are estimated from the sample of data. Common practice dictates that if $E_i < 5$, then cell i is combined with a neighboring cell. Also, for a closer approximation to the chi-square distribution, G is commonly adjusted as follows:

$$G_{adj} = G\ /\ \{1 + (k^2-1)/[6N(k-p-1)]\} \qquad (3.5\text{-}2)$$

where N is the sample size.

Unlike the preferred G test, the basis for the traditional chi-square goodness-of-fit test is more intuitive than theoretical. Its test statistic is a measure of the difference between observed and expected cell frequencies, squared to get positive differences, expressed as proportions of the expected frequencies, and summed over all cells:

$$X^2 = (O_1 - E_1)^2/E_1 + \ldots + (O_k - E_k)^2/E_k \qquad (3.5\text{-}3)$$

Like G, X^2 is approximately distributed as a chi-square variate with k-1-p degrees of freedom. Also, as in the G test, if $E_i < 5$, then cell i is commonly combined with a neighboring cell.

Both the G and chi-square tests for goodness-of-fit have several advantages over the standard K-S test:

- they work well with discrete data,
- simple adjustments are available for intrinsic null hypotheses, and
- they are not as sensitive to serial correlation.

The K-S test would clearly be the preferred test if the extrinsic and continuity assumptions were valid. Since they are not, the G test is the preferred test. The K-S test will be performed because of its useful interpretation as a maximum error bound for the AFGL models. The chi-square test will be performed because it is the traditional goodness-of-fit test and because its computation will not appreciably impact project cost.

3.6 SIMULATION STUDIES

As stated in Section 3.3, the criterion for selecting the model validation sample size is that it should be the smallest

integer multiple of 150 observations, (i.e., 150, 300, 450, ...) which will give us 95 percent confidence in rejecting a null hypothesis of equality of empirical and modeled distributions if the absolute value of the difference between the corresponding CDFs anywhere exceeds 0.05. There is a natural and close relationship among this sample size criterion, the AFGL requirement to determine model error bounds, and the definition of the K-S statistic.

While it would be desirable to employ a K-S statistic in determining model error bounds, it is not possible to use published tables or K-S critical values for reasons discussed above. A standard procedure in such situations is to use Monte Carlo methods to randomly generate empirical distributions of K-S statistics with the desired properties. In this case the properties are:

- discrete, serially correlated data
- intrinsic null hypotheses
- geometries specified by the structure and resolution of GOES satellite imagery.

Once relevant K-S distributions are obtained, interpolation provides the desired critical values (i.e., the 95th percentiles).

As shown above, it was not absolutely necessary to determine specific K-S critical values to obtain useful estimates of required sample sizes. However, everyone agreed that doing so would be useful. The Monte Carlo estimation procedure entails simulation of most of the model validation procedure and thus can be viewed as a practice run. Credit for the original proposal to use Monte Carlo simulation to estimate K-S critical values goes to Major Albert Boehm of AFGL.

To reduce computer costs, several constraints were imposed on the simulations:

- only the line algorithm (BLA) would be used
- only two extreme pairs of the model parameters (p0 and r) would be used
- only three sample sizes would be used: 30, 150, and 450. K-S critical values for other sample sizes would be estimated by curve fitting.

As it turned out, the first constraint was relaxed somewhat and one simulation experiment was performed with the area algorithm (BAA).

High level pseudocode for the Monte Carlo simulation procedure is shown in Fig. 3.6-1. The first step was to determine two extreme, but feasible pairs of the model parameters:

p0 and r. As discussed in Section 2.2-1, the three sites chosen for the model validation are Ft. Riley, KS; Rickenbacker AFB, OH; and Key West, FL. At each of these sites, the parameters p0 and r vary as a function of time of day and season. It is likely that the parameters also vary with other factors such as synoptic situation, but no other factors were considered. Figure 3.6-2 snows a two dimensional scatterplot of $|p0-0.5|$ vs r for three RUSSWO sites which are very close to the model validation sites: Wichita (Ft. Riley), Columbus (Rickenbacker AFB), and Key West. Each point represents a pair of parameter values for one of three sites, one season, and one of four times a day.

The polygon shown in Fig. 3.6-2 is a convex hull for the sample point set. That is, it is a convex polygon with the smallest area such that each point in the sample point set falls on or within the polygon. The convex hull bounds the sample feasible region (F) of the parameter space. A well-known theorem of linear programming states that extreme

```
SELECT TWO EXTREME SITES (p0, r)
FOR SITE = 1 TO 2 DO
    FOR REPETITION = 1 TO 100 DO
        FOR SAMPLE SIZE = 30, 150, AND 450 DO
            FOR LINE LENGTH = 10, 50, AND 300 KM DO
                GENERATE SYNTHETIC IMAGES
                BUILD FRACTIONAL COVER CDFs
            END FOR
            USE 50 KM CDF TO ESTIMATE (p̂0, r̂)
            FOR LINE LENGTH = 10, 50, AND 300 KM DO
                COMPUTE THEORETICAL CDF USING BLA
                COMPUTE K-S STATISITICS
            END FOR
        END FOR
    END FOR
    BUILD CDFs FOR K-S STATISTICS
    INTERPOLATE TO 95th PERCENTILES
    MODEL 95th PERCENTILES AS FUNCTIONS OF SAMPLE SIZE
    INTERPOLATE TO SAMPLE SIZES REQUIRED FOR K-S ≤0.05
END FOR
```

Figure 3.6-1 Monte Carlo Simulation Procedure

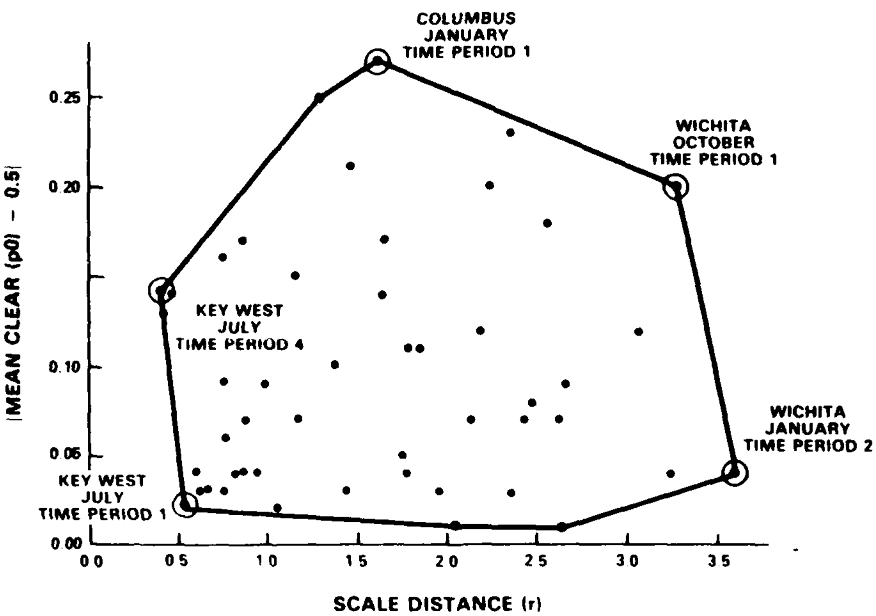


Figure 3.6-2 Wichita, Columbus, and Key West Sample Parameter Space

values of any linear function defined on F occur at vertices of the convex hull bounding F. In that sense, each of the eight vertices in Fig. 3.6-2 is an extreme point representing an extreme pair of sample parameter values. Five prominent extreme points have been circled and identified. For the Monte Carlo simulation, we chose to use the extreme points farthest from and closest to the origin of the scatterplot -- Site 1: Wichita, October, Time Period 1 (p0 = 0.30 r = 3.29) and Site 2: Key West, July, Time Period 1 (p0 = 0.48 r = 0.55).

Synthetic satellite images were generated using Boehm's three-dimensional *Sawtooth Wave Model*. Model parameters were selected so that:

- the first two dimensions were spatial and the third was temporal
- images for sequential days were correlated with a correlation coefficient of 0.23 (this fixed the temporal wavelength)
- the horizontal wavelength was fixed by scale distance (3.29 or 0.55)
- the geometry of a GOES scene was idealized as 300 x 300 array of 1 km square pixels.

The simulation procedure outlined in Fig. 3.6-1 involved generating samples of synthetic images, analyzing each image to determine the fractional sky cover along 10, 50, and 300 km lines, and then building a sky cover sample CDF. Next, the sample CDF for the 50 km line was used to estimate the sample parameters (p0 and r). These parameters were then fed into the BLA to obtain theoretical (model predicted) CDFs for the 10, 50, and 300 km lines. Comparing the sample CDFs with the theoretical CDFs resulted in one K-S statistic for

each length line. This entire process was repeated 100 times to obtain empirical distributions of K-S statistics for each length line. The K-S critical values, defined to be the 95th percentiles of the K-S distributions, were obtained by interpolating within the empirical distributions. The results of the simulation process up to this point are summarized in Table 3.6-1.

Examining Table 3.6-1, we immediately note that the length of the longest line at Site 2 is 140 km rather than 300 km. This is because the BLA model is valid for lengths between 0.5r and 256r. At Site 2, r = 0.55 and the longest valid line is 140.8 km. The BAA model has the same length scale limitation. For situations like this one where the scale distance is small, the BLA and BAA models, in their current form, are limited to applications with relatively short lines and small areas.

TABLE 3.6-1

K-S CRITICAL VALUES FROM MONTE CARLO SIMULATION (BLA, 100 REPETITIONS, 50 Km LINE CDF USED TO ESTIMATE PARAMETERS)

Sample Size	Site 1			Site 2		
	10 Km	50 Km	300 Km	10 Km	50 Km	140 Km
30	.0538	.0650	.2895	.1179	.1506	.2163
150	.0218	.0304	.1382	.0565	.0561	.0999
450	.0126	.0188	.0724	.0366	.0395	.0540

Examining Table 3.6-1 further, we note:

- critical values decrease as sample size increases
- critical values generally increase as line length increases
- a sample size of 450 would not be adequate for the longest lines if the model parameters were estimated from the CDF for the 50 km line.

The last observation prompted additional simulation experiments which will be described shortly.

In Section 3.4, it was pointed out that for many modified K-S statistics which have been developed using Monte Carlo simulation, as well as for the standard K-S statistic, the large sample asymptotic relation between critical value and sample size has the general form:

$$D_{0.05} = \text{constant}\ N^{-0.5} \qquad (3.6\text{-}1)$$

Equation 3.6-1 states that the reciprocal of the K-S critical value is proportional to the square root of the sample size.

To verify that Eq. 3.6-1 applies to the data in Table 3.6-1, we may use the known critical values and sample sizes and compute the "constant" values. The results are presented in Table 3.6-2.

Note that there is very little variation in the estimated constant values as a function of sample size. This is good news. It means that Monte Carlo simulation with a sample size as small as N = 30 can produce reasonably accurate estimates of the required sample sizes.

TABLE 3.6-2

"CONSTANTS" FROM EQ. (3.6-1) APPLIED TO DATA IN TABLE 3.6-1

Sample Size	Site 1 10 km	Site 1 50 km	Site 1 300 km	Site 2 10 km	Site 2 50 km	Site 2 140 km
30	0.29	0.36	1.59	0.65	0.83	1.18
150	0.27	0.37	1.69	0.69	0.69	1.22
450	0.27	0.40	1.54	0.78	0.84	1.15

To estimate the required sample sizes, we apply Eq. (3.6-1) a second time using $D_{0.05} = 0.05$ and the constant values from the last row (N = 450) of Table 3.6-2. The results are presented in Table 3.6-3.

As noted above, when the CDF from the 50 km line is used to estimate parameters, a sample larger than 450 will be required to determine error bounds when the BLA model is applied to the longest lines (300 and 140 km here).

The natural question is: How will these results change if the CDFs from other length lines are used to estimate the parameters? To find out, the Monte Carlo

TABLE 3.6-3

REQUIRED SAMPLE SIZES FROM MONTE CARLO SIMULATION (BLA, N = 450, 100 REPETITIONS, 50 Km LINE USED TO ESTIMATE PARAMETERS)

Site 1 10 Km	Site 1 50 Km	Site 1 300 Km	Site 2 10 Km	Site 2 50 Km	Site 2 140 Km
29	64	944	242	281	525

simulation procedure specified in Figure 3.6-1 was repeated with the following modifications:

- sample size was fixed at N=30
- site was not varied (Site 1 was used as it required the largest sample in the previous simulations)
- parameters were estimated from the CDF for each line length and K-S critical values were determined for each set of parameter values
- required samples sizes were estimated using Eq. (3.6-1) twice, as described above.

The results are presented in Table 3.-6-4.

Table 3.6-4 is enlightening. Examining the columns, we observe that the smallest required sample size occurs when the CDF for the line length under consideration is used to estimate the parameters. In application terms, this implies that the BLA and BAA will be most accurate when they are used to smooth known empirical CDFs. They will be less accurate when they are used to extrapolate from known CDFs to CDFs for other lengths and areas. Intuitively, this result is expected.

TABLE 3.6-4

K-S CRITICAL VALUES AND REQUIRED SAMPLE SIZES FROM MONTE CARLO SIMULATION (SITE 1, BLA, N =30, 100 REPETITIONS)

Parameter Estimation Line Length (Km)	K-S Critical Values 10 Km	50 Km	300 Km	Required Sample Size 10 Km	50 Km	300 Km
10	.0375	.1691	.3230	17	343	1252
50	.0538	.0650	.2895	35	51	1006
300	.1076	.1127	.1324	139	152	211

Suppose we wish to estimate model parameters once, and apply the models with various lines and areas. Examining the rows of Table 3.6-4, we observe that the best strategy is to estimate the parameters using the largest available area or longest available line. We can expect the BLA and BAA models then to be reasonably accurate if they are applied on shorter lines and over smaller areas. Using the models to extrapolate CDFs to shorter lines and smaller areas is a stable process more akin to interpolation than to extrapolation. Conversely, using the models to extrapolate CDFs to larger lines and areas is an unstable process characterized by more variability and by greater errors.

With respect to required sample size, we conclude that a sample of 150 is not adequate. A sample of 300 or 450 should be adequate, providing that the CDF from the longest line or preferably the largest area is used in the parameter estimation.

Although beyond the scope of the original plans, one Monte Carlo experiment with the BAA algorithm was performed to see if the results would be much different. The procedure specified in Fig. 3.6-1 was employed with the following modifications:

- site was fixed (site 2 was used as its largest area was only $(140\ \text{km})^2$ and computer cost increased with area
- ten repetitions were used rather than 100 (again to control computer costs)
- sample size was fixed at N = 30
- areas were $(10\ \text{km})^2$, $(50\ \text{km})^2$, and $(140\ \text{km})^2$
- the BAA model was employed.

Note that the CDF from the $(50\ \text{km})^2$ area was used to estimate parameters. The results are presented in Table 3.6-5.

Comparing the required sample sizes in Table 3.6-5 to the comparable quantities in Table 3.6-3, we observe that extrapolating to larger areas is even more damaging than extrapolating to longer lines. This result reinforces our conclusion that in determing K-S critical values, the CDF from the largest area should be used to estimate the model parameters.

In summary, we have shown that for the two sites considered, a sample size of 300 or 450 will suffice for reliable estimation of BLA model error bounds. Considering the focus on just two extreme points in Figure 3.6-2, it seems prudent to recommend a sample size of 450 rather than 300. Comparison of the very limited results from the BAA simulation with comparable results from the BLA simulation adds additional impetus for recommending a sample size of 450. These conclusions will not be valid if either model is used to extrapolate a sky cover CDF to a larger area or longer line. Consequently, it is imperative that the CDF from the largest area or longest line be used when estimating model parameters.

TABLE 3.6-5

K-S CRITICAL VALUES AND REQUIRED SAMPLE SIZES FROM MONTE CARLO SIMULATION (BAA, 10 REPETITIONS, N = 30, $(50\ \text{Km})^2$ AREA USED TO ESTIMATE PARAMERERS)

K-S Critical Values			Required Sample Sizes		
$(10\ \text{Km})^2$	$(50\ \text{Km})^2$	$(140\ \text{Km})^2$	$(10\ \text{Km})^2$	$(50\ \text{Km})^2$	$(140\ \text{Km})^2$
.1362	.1272	.2884	223	195	998

4. SUMMARY AND FUTURE PLANS

The purpose of this project is to determine error bounds for two empirical models of sky cover climatology: the BLA and BAA models. To accomplish this objective, there are six study tasks:

- Task 1 - Select Test Areas
- Task 2 - Determine Sample Sizes
- Task 3 - Acquire Satellite Imagery
- Task 4 - Develop Cloud Detection and Analysis Software
- Task 5 - Create a Cloud Cover Database
- Task 6 - Determine Error Bounds for the AFGL Models.

During the period of this report the first four tasks have been completed.

Three sites (centered on Ft. Riley, Rickenbacker AFB, and Key West) were selected. Selection criteria included: obtaining cloud cover distributions of a variety of shapes, homogeneous climatology in the vicinity of each site, proximity to the GOES subpoint, proximity to a RUSSWO site, and availability of concurrent surface data and satellite imagery.

Required sample sizes were determined to be 450 scenes for each location, season, and time of day. Sophisticated Monte Carlo simulation of the model validation procedure provided other valuable insights and guidelines. These include:

- using the models to extrapolate existing CDFs to larger lines and areas is an unstable process and can be expected to be error prone

- using the models to smooth existing CDFs or to extrapolate to smaller areas and lines is a stable process and can be expected to have manageable error properties

- during the determination of model error bounds, model parameters should be estimated from CDFs for the largest area, or longest line, or both. Optionally, they may also be estimated from CDFs for other lines and areas.

GOES imagery for three sites, two times of day, and two seasons were acquired. Included are a total of 5460 bispectral (visible and IR) scenes (316 Km square).

Automated cloud detection and analysis algorithms have been developed, coded, and tested. A quality control feature identifies suspect images for manual interactive evaluation.

Construction of the cloud cover database is now underway. Scenes will be analyzed and climatologies constructed in station order sequence so that AFGL can commence their task of determining empirical K-S statistic distributions. Once relevant K-S critical values have been determined, RDS with assistance from TASC, will determine model error bounds using three goodness-of-fit tests (log-likelihood ration, chi-square and K-S).

APPENDIX A
IMAGE PROCESSING ALGORITHMS

There are three primary programs which have been written to extract cloud and surface data and process this information into the cloud detection data base. These programs are described below.

A.1 DATSAV TAPE READING PROGRAM

The user must first run the DATSAV tape reading program (Figure A.1) which will extract data from tape files to files which can then be used by the automatic cloud detection program. A brief description follows:

1) User runs the DATSAV tape read program to extract data from the DATSAV tapes for the particular station, year, month, and hours.

2) The program produces a file consisting of surface data for a particular station, year, month, and two hours per day.

A.2 THE MASTER PROGRAM

A cloud detection algorithm has been formulated and is in the final stages of testing. The program uses the GOES images over FL, OH, and KS and the corresponding DATSAV data files for surface observations (ground truth). The following summarizes the steps taken by the MASTER program. (See Figure A.2.)

1) Read the images off the GOES tapes for the appropriate 316 km x 316 km subscene (corresponding to a 100,000 square kilometer area) centered on each of the three ground stations of interest.

2) Register the IR and VIS scenes to one another.

3) Compute a minimum background intensity level map based on five days of images and used for each of those five days (VIS and IR).

4) Locate the mode (peak value) in the background map, and set the entire background map equal to this constant value (VIS and IR).

5) Subtract this background from the image for the day of interest (VIS and IR).

6) Threshold the image (VIS and IR) result at a low residual level (10 to 15) to correct for noise in the image, depending on the station.

7) Compute the cloud cover over 50 km x 50 km regions centered on the ground station for each scene (VIS and IR).

8) Reject the image if the IR cloud determination is greater than three deciles (30%) above the VIS cloud determination.

9) Reject the image if the DATSAV (ground truth) sky cover value differs by more than twenty-five percent from the VIS cloud determination.

10) If not rejected, compute cloud cover over all reference lines and areas, and maximum clear and cloudy runs over all reference lines. Write image file name and image descriptive data to the "GOOD" image data file and update the "CUMULATIVE" data file with the various bin counts.

11) If rejected, write the image file name and associated image descriptive data to the "BAD" image data file.

A.3 THE MASTERI PROGRAM

The third and final program, MASTERI, (Figure A.3) interactively selects the bad images which the user wishes to display, examine, and interactively edit for possible inclusion in the valid image data base. The bad images are listed within the "BAD" image file. This process consists of the following steps:

1) Displaying a selected image on the screen.

2) Interactively threshold the visible image through inspection of the displayed image.

3) Give the interactive program the selected threshold value for computation of new cloud cover statistics.

4) The program updates the "GOOD" and "CUMULATIVE" data files.

5) User can select another image to view if desired.

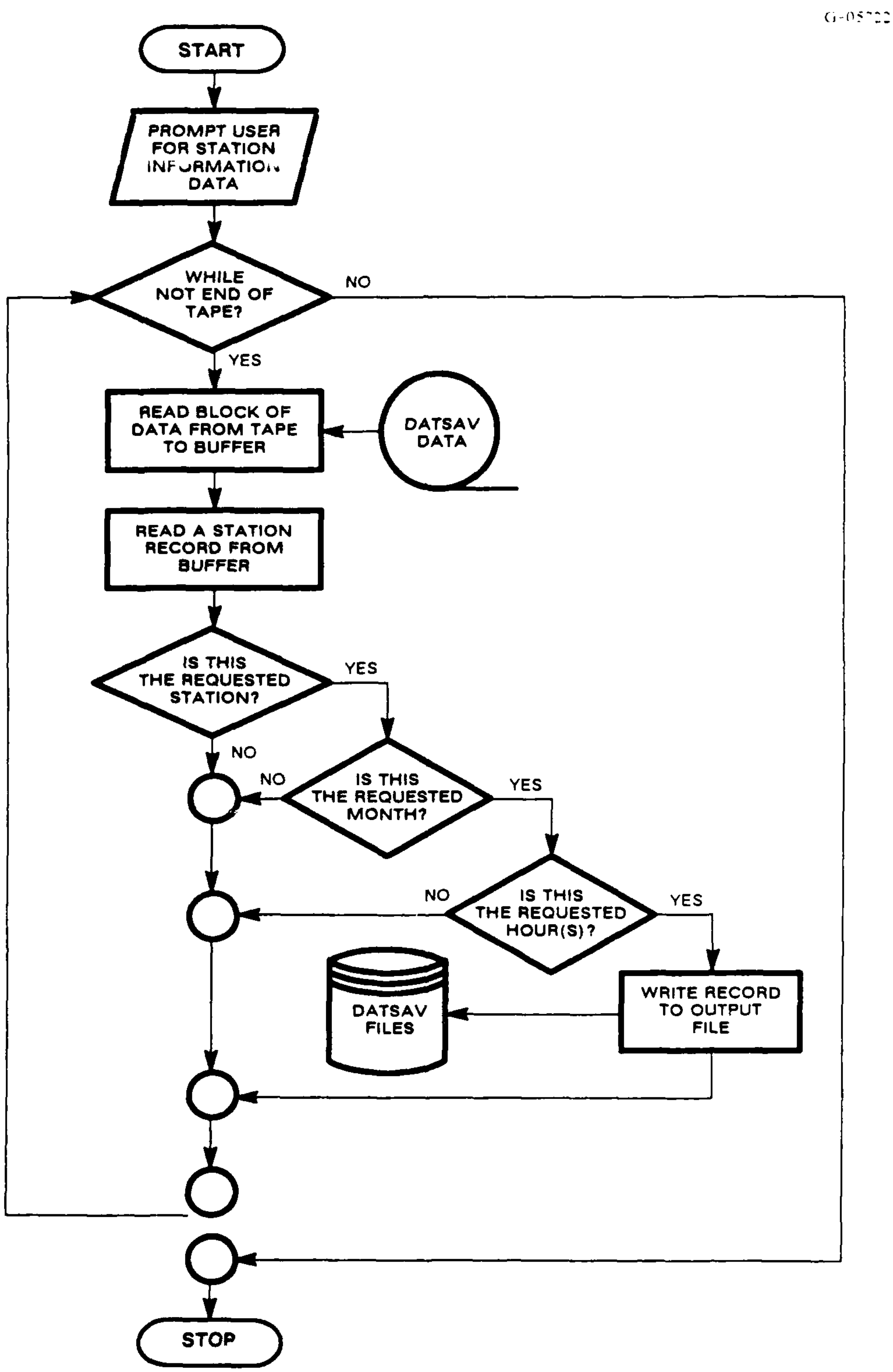


Figure A.1 DATSAV Tape Reading Program Flow Diagram

G-05718

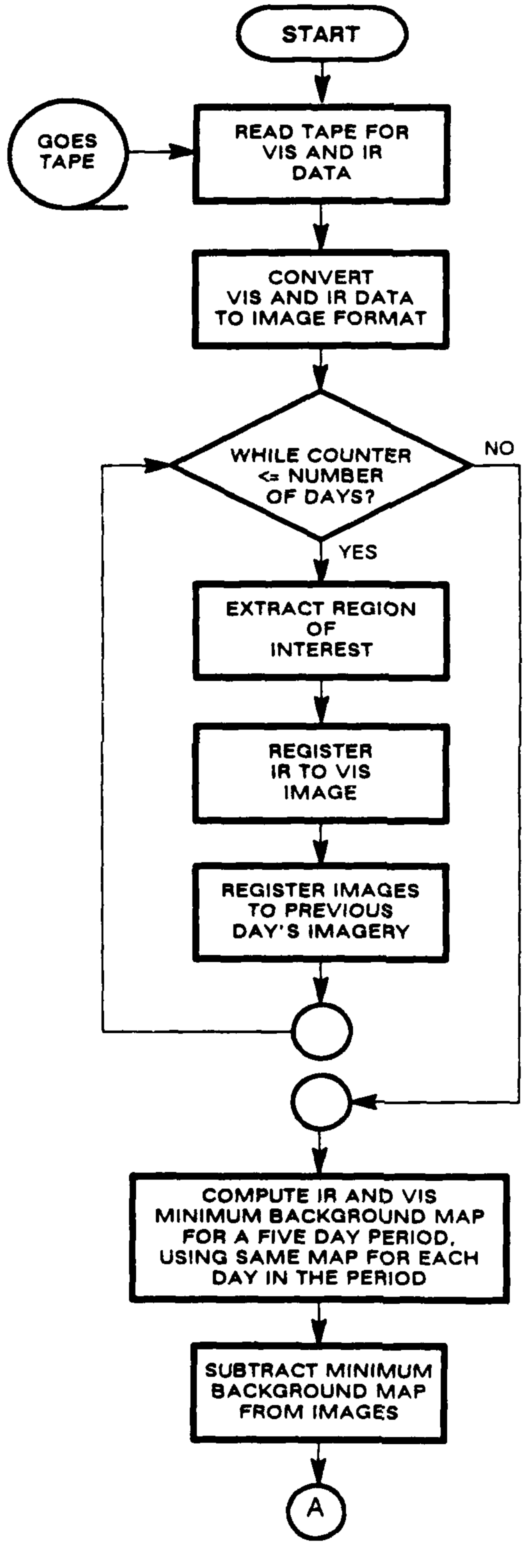


Figure A.2 MASTER Program Flow Diagram

G-05~19

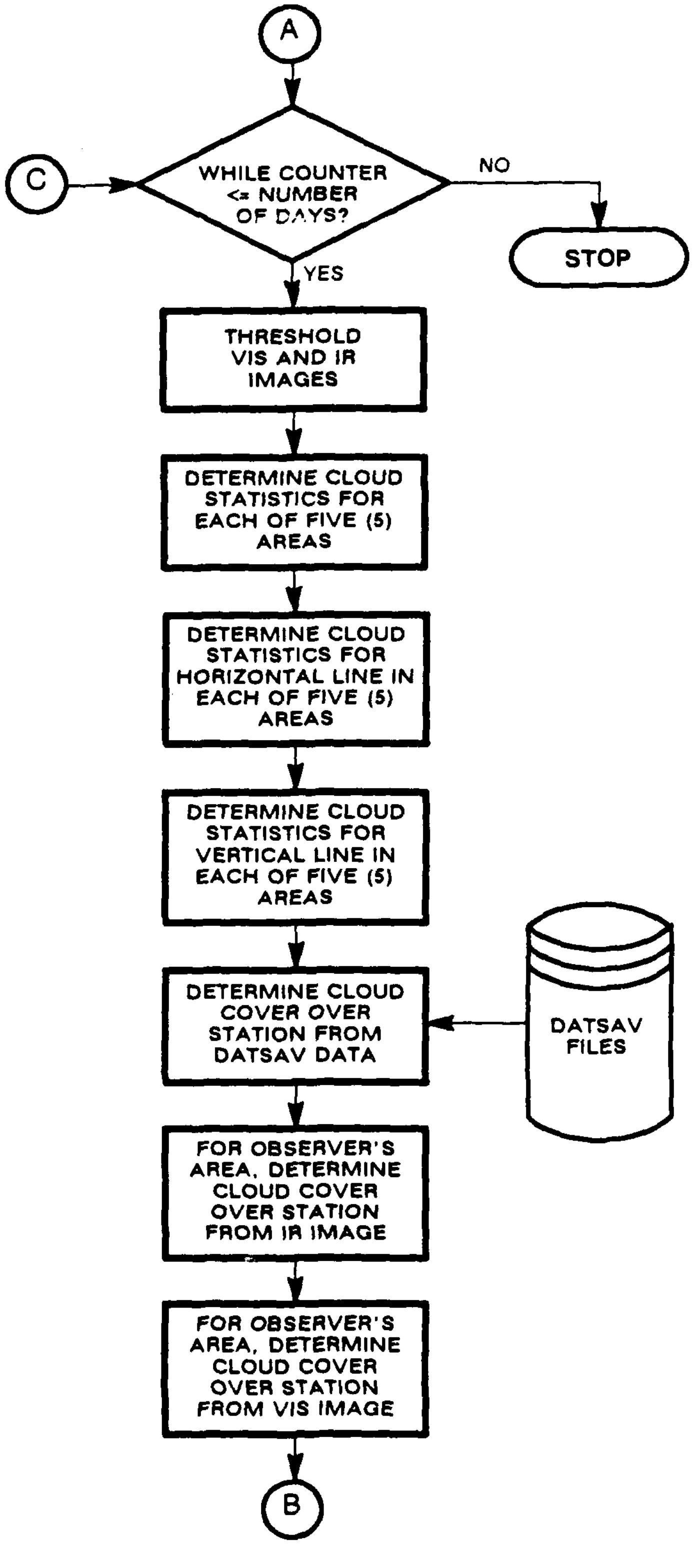


Figure A.2 MASTER Program Flow Diagram (Cont'd)

G-05720

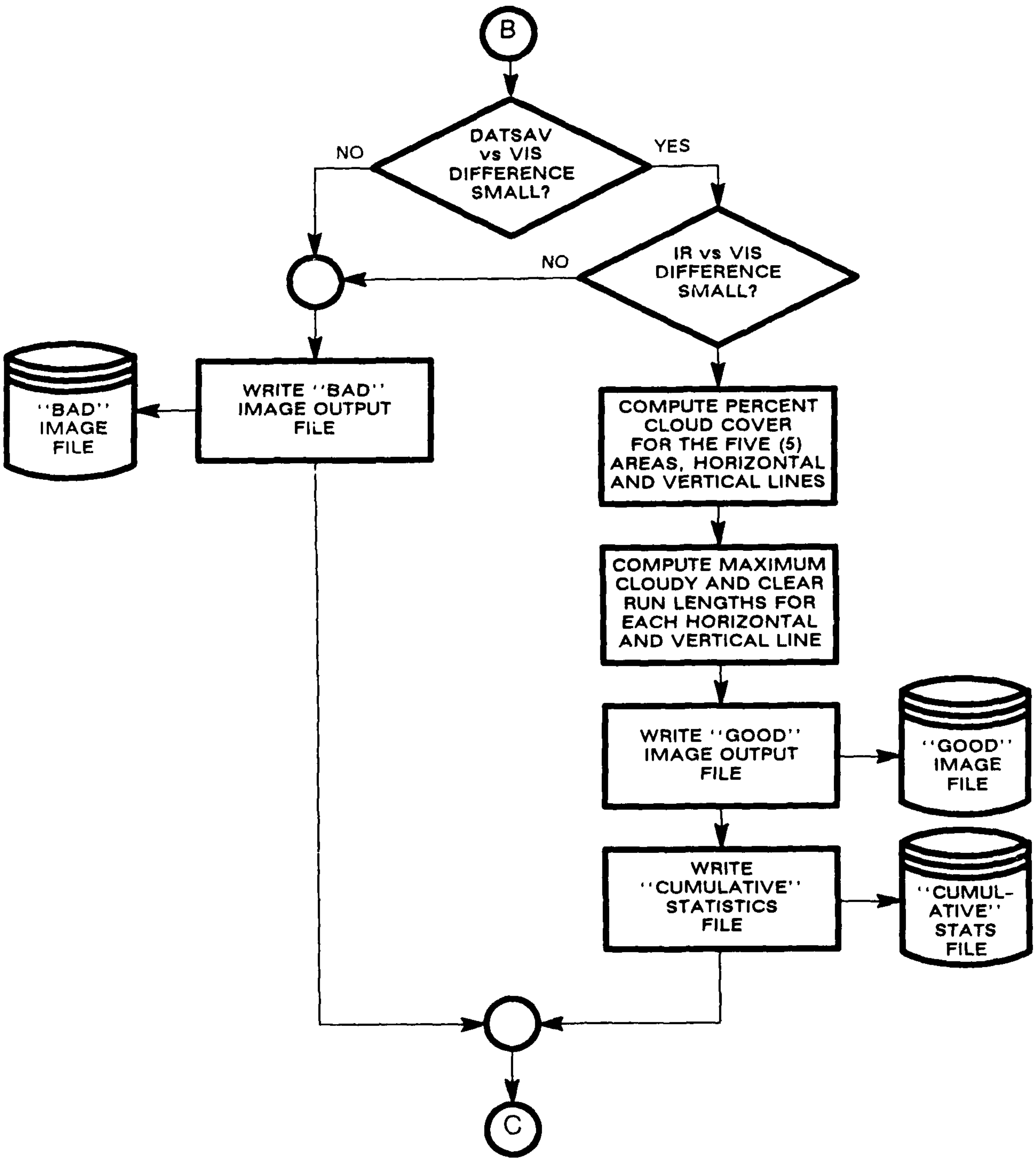


Figure A.2 MASTER Program Flow Diagram (Cont'd)

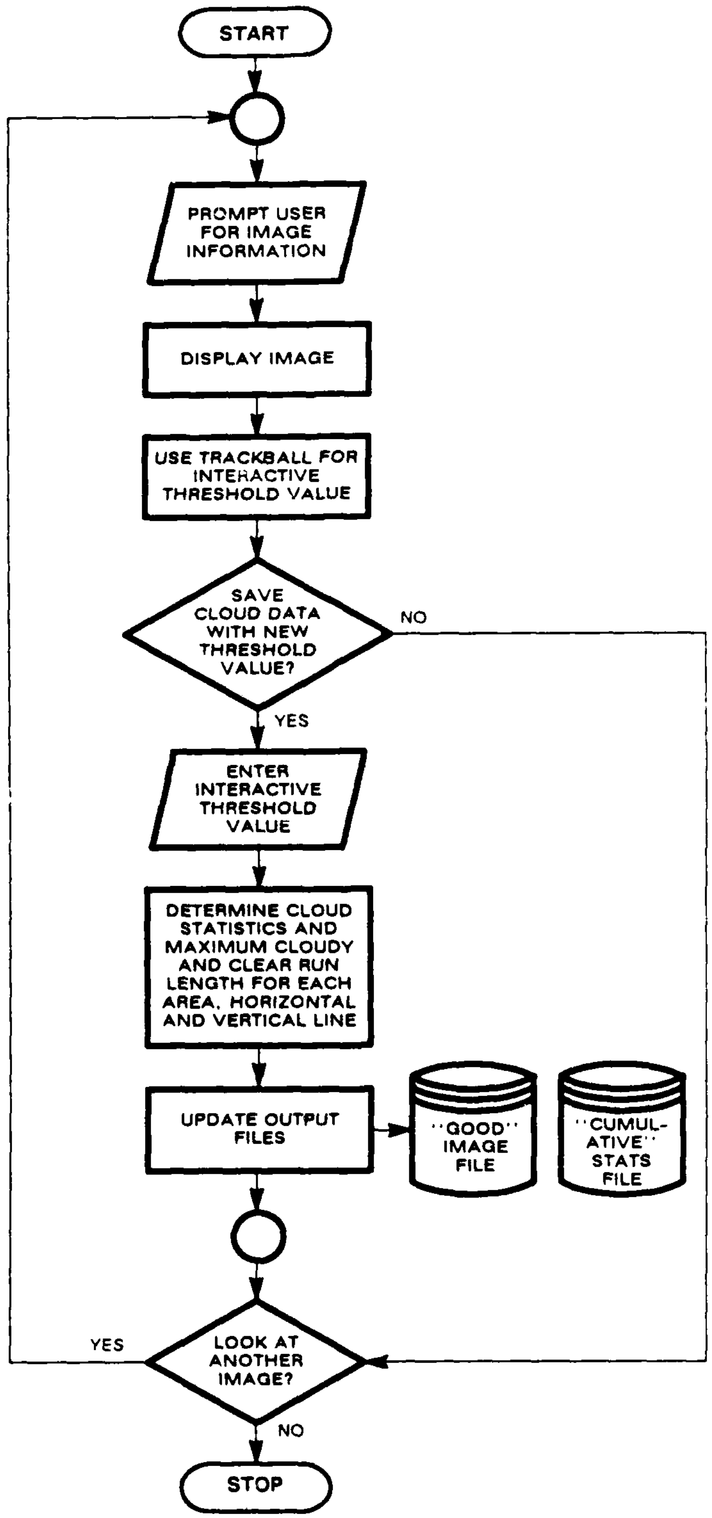


Figure A.3 MASTERI Program Flow Diagram

APPENDIX B

OUTPUT FILES

CONTENTS OF OUTPUT STATISTICS FILE

Description	Data Type	Range
Image Name (hour, month, day, VIS or IR scene)	Char. String	-
Cloud cover over station, computed (decile)	Integer (nnn)	0-10
Cloud cover over station, computed (fraction)	Float (n.nnnn)	0.0-1.0
Cloud cover over station, ground truth from DATSAV data (Synoptic or Airways)	Integer (nnn)	0-8 or -2,-7,-8,-9
Cloud cover over station, computed from IR Image (decile)	Integer (nnn)	0-10
Background level-visible threshold	Integer (nnnn)	0-255
For each of 5 areas*: cloud cover over area, computed (fraction)	Float (n.nnnnn)	0.0-1.0
For each of 5 Horizontal lines+: cloud cover along line, computed (fraction)	Float (n.nnnnn)	0.0-316.0
Maximum clear run length (km)	Float (nnn.nnnnnn)	0.0-316.0
Maximum cloudy run length (km)	Float (nnn.nnnnn)	0.0-316.0
For each of 5 Vertical lines:++ (same as for Horizontal lines)		

*Areas have sides of 10, 32, 50, 100, and 316 km.
+Lines have the same length as the areas.
++Lines have the same height as the areas.

Note: There will be a separate output statistics file for each of 3 scenes, 2 times of day, 2 seasons, 5 years = 60 files.

Each file has a record for each day for 1 season (3 months).

CONTENTS OF OUTPUT CUMULATIVE STATISTICS FILE

Description	Data Type	Range
For each of 5 Areas:		
For each of 22 Bins:		
Number of days with cloud cover in this bin	Integer (nnnnnn)	0-999999
For each of 5 Horizontal Lines:		
For each of 22 Bins:		
Number of days with cloud cover in this bin	Integer (nnnnnn)	0-999999
For each of 317 Bins:		
Number of days with maximum length clear run in this bin (KM)	Integer (nnnnnn)	0-999999
Number of days with maximum length cloudy run in this bin (KM)	Integer (nnnnnn)	0-999999
For each of 5 Vertical Lines:		
(same as horizontal lines)		

REFERENCES

1. Burger, C.F., and I.I. Gringorten, 1984: Two-Dimensional Modeling for Lineal and Areal Probabilities of Weather Conditions. Environmental Research Papers No. 875, AFGL-TR-84-0126. U.S. Air Force Geophysics Laboratory, Hanscom AFB, MA 01731. 58 pp. ADA147970

2. Snedecor, George W. and William G. Cochran, 1967: Statistical Methods (Sixth Edition). Iowa State University Press, Ames. 593 pp.

3. Lapin, Lawrence, 1978: Statistics for Modern Business Decisions. Harcourt, Brace, Jovanovich Inc., New York. 856 pp.

4. Rohlf, F. James and Robert R. Sokal, 1981: Statistical Tables. W. H. Freeman, San Francisco. 219 pp.

5. Sokal, Robert R. and F. James Rohlf, 1981: Biometry. W. H. Freeman, San Francisco. 859 pp.

6. Lilliefors, H. W., 1967: The Kolmogorov-Smirnov test for normality with mean and variance unknown. J. Amer. Stat. Associ., 62, pp 399-402.

7. Crutcher, Harold L., 1975: A note on the possible misuse of the Kolmogorov-Smirnov test. J. Appl. Meteor., 14, pp 1600-1603.

8. Gleser, L. J., 1985: Exact power of goodness-of-fit tests of Kolmogorov type for discontinuous distributions. J. Amer. Stat. Assoc., 80, pp 954-958.

9. Pettitt, A. N., and M. A. Stephens, 1977: The Kolmogorov-Smirnov goodness-of-fit statistic with discrete and grouped data. Technometrics, 19, pp 205-210.

10. McCabe, John T., 1968: Estimating Conditional Probability and Persistence. Air Weather Service Technical Report 208, U.S. Air Force Weather Service. 17 pp.